



Systemic evaluation governance. New logics in the Development of organsational fields

Hansen, Hanne Foss

Published in:
Scandinavian Journal of Public Administration

Publication date:
2013

Citation for published version (APA):
Hansen, H. F. (2013). Systemic evaluation governance. New logics in the Development of organsational fields. *Scandinavian Journal of Public Administration*, 16(3), 47-64.

Hanne Foss Hansen
Institut for Statskundskab,
Københavns Universitet

Keywords: Evaluation,
governance, performance,
evidence, organisational
fields

Nøgleord: Evaluering,
styring, performance, evi-
dens, organisationsfelter

Abstract

Systemic evaluation governance, defined as governance systems anchored in evaluation and focused on developing performance through comparing actors in organisational fields, is increasingly institutionalised at international as well as national levels in recent years. Systemic evaluation governance takes several forms, e.g., the forms of indicator systems, benchmarking, accreditation, certification and initiatives of evidence-based professional practice. On the basis on a discussion of the concept of systemic evaluation governance, the article develops a typology of different forms of this specific governance type and analyses and discusses the logics and premises these are anchored in. The analysis is illustrated by empirical examples.

Systemisk evalueringsstyring. Nye logikker i udvikling af organisationsfelter

Systemisk evalueringsstyring, defineret som styringssystemer der er forankret i evaluering og har fokus på at skabe forbedrede resultater gennem sammenligning af aktører i et organisationsfelt, bliver i stigende grad institutionaliseret internationalt såvel som nationalt i disse år. Systemisk evalueringsstyring findes i flere varianter, f. eks. i form af indikatorer, benchmarking, akkreditering, certificering og evidensbaseret af fagprofessionel praksis. Med afsæt i en diskussion af begrebet systemisk evalueringsstyring, udvikles en typologi over de forskellige varianter og det diskuteres hvilke logikker og præmisser disse bygger på. Analysen illustreres ved empiriske eksempler.

Offentlig förvaltning
Scandinavian Journal of
Public Administration
16(3): 47-64

© Hanne Foss Hansen and
Förvaltningshögskolan, 2013
ISSN: 1402-8700
e-ISSN: 2001-3310

***Hanne Foss Hansen**, ph.d., er professor i offentlig forvaltning og organisation på Institut for Statskundskab på Københavns Universitet. Hun er forfatter til flere bøger og artikler om offentlig organisation, reformer, forvaltningspolitik, evaluering og evidens. Hun er for øjeblikket ansvarlig for et forskningsprojekt vedrørende krisens forvaltningspolitiske indflydelse.

Introduction

Evaluation is practiced in a variety of ways. New methodologies are being developed and ideas about why and how to evaluate change across time. In recent years, evaluation has been increasingly built into comprehensive governance initiatives. Evaluation governance has become increasingly popular as it spreads across countries and policy fields. As an effect, competition and control in society are enhanced among actors at many levels, from the international and European arenas to local government. This article addresses this phenomenon. By suggesting and developing the concept of systemic evaluation governance (SEG), these evaluation governance systems are analysed. SEG is defined as evaluation carried out with steering ambitions and targeted at several actors in a field. SEG for example may be targeted at several organisations in an organisational field, e.g., a policy sector, as is the case when educational institutions are benchmarked aiming at supporting free user choice, or it may be targeted at several nation states, as is the case in relation to the Open Coordination Method at the European Union level.

The concept of SEG will be more thoroughly discussed below. As yet, it is enough to emphasise that systemic evaluation governance comes in various forms, such as benchmarking, ranking, rating, accreditation, certification as well as systemic initiatives to spread knowledge and secure evidence based policies and practices. At national levels, SEG is developed so to speak on top of already existing hierarchical performance management systems. At the EU level, SEG is considered an alternative to classical rule-based government.

The main research questions in this article concerns the following. What are the characteristics of SEG conceptually and in practice? And what are the logics of the different forms of SEG? In order to answer these questions, the article presents a conceptual and descriptive analysis of SEG. The aim is to contribute to our understanding of SEG as a phenomenon and to determine important analytical dimensions, which can be used for more in-depth empirical analyses of SEG and its consequences.

The article is organised in six sections. Section two presents the conceptual discussion of SEG. Section three reflects on SEG in a historical perspective. Section four presents a typology of different forms of SEG and discusses the logics in which they are anchored. In addition, it presents examples of different forms of SEG from the Danish public sector. Section five looks ahead and reflects upon which direction the development of SEG is heading in the coming years. Finally, section six provides the conclusion.

Central concepts

By way of introduction, the content of the concept of SEG will be defined. This will be done by discussing the concepts of evaluation and governance as well as by clarifying what is meant by systemic.

In common language, evaluation simply means assessment. In the specialised literature on evaluation, the content is developed further. We find different

definitions underlining slightly different aspects. Here are two samples. Evaluation is:

“...an applied inquiry process for collecting and synthesising evidence that culminates in conclusions about the state of affairs, value, merit, worth, significance, or quality of a program, product, person, policy proposal or plan. Conclusions made in evaluations encompass both an empirical aspect (that something is the case) and a normative aspect (judgement of the value of something).” (Fournier, 2005: 140).

“...careful retrospective assessment of the merit, worth and value of administration, output, and outcome of government interventions, which is intended to play a role in future, practical action situations.” (Vedung, 1997: 3).

Together, the two definitions indicate that evaluation may include several evaluation criteria (“value, merit, worth, significance, or quality”) and focus on several evaluands (“program, product, person, policy proposal or plan”, “administration, output, and outcome of government interventions”). In addition, they tell us that evaluation demands data collection and data analysis using systematic methodology as a basis for assessment.

The two definitions, however, originate from different evaluation traditions. Whereas the first one reflects the American traditions for programme evaluation, the second one reflects the Scandinavian welfare state context and thus explicitly includes an organisational focus.

The concept of governance is elastic as well. Below the concept of governance is used as a concept broader than the classic concept of government but at the same time as a concept more narrow than the definition suggested by Pierre (2000), who defines governance broadly as coordination and coherence among a wide variety of actors with different purposes. To be more specific, governance is defined below as having to do with implementation of policies through systems and instruments sometimes presented as governance modes (Howlett, 2009: 76). The literature on governance modes has been focused on developing typologies. Jørgensen and Larsen (1982) suggest a distinction between bureaucracy, market, democracy, norms and knowledge. Vedung (1998) suggests a distinction between rules (also termed “sticks”), economic resources (also termed “carrots”) and information (also termed “sermon”), and Howlett (2009) suggests a distinction between legal, corporatist, market and network governance.

There are similarities as well as differences across these typologies. The categories of bureaucracy, rules and legal, market and economic resources, as well as knowledge and information, are clearly related. While norms belong in a separate category in the first typology, norms are defined as smooth rules in the second and not explicitly included in the third. There are two very important differences between the typologies. The first concerns the concept of democracy. The inclusion of democracy in the first typology can be interpreted as an empha-

sis on the premise that democracy may be implemented not only as a political but also as an organisational governance system. On the contrary, the exclusion of democracy from the second and third typologies probably is an expression of the view that democracy is understood as an overall frame for political decisions concerning which governance modes to use in specific policy fields. The second important difference is that while typologies one and two focus primarily on governance instruments, typology three explicitly includes institutional aspects in the modes of corporatist governance aiming at creating development through the management of major organised social actors and network governance, with the purpose of development through co-optation and self-organisation of social actors through voluntarily collaboration.

Governance may fulfil several goals. First and foremost, as mentioned above, governance aims at implementing and fulfilling politically decided policies and values, such as quality, economic responsibility, the rule of law or environmental sustainability. Secondly, governance aims at securing accountability and legitimacy, for example by securing that organisations document results and thereby show they are “accountable”.

Having discussed the concepts of evaluation and governance, we now examine the concept of evaluation governance. Evaluation governance refers to governance based directly on evaluative information, that is governance initiatives anchored in systematic assessment of organisation, implementation, output and outcomes of public policy. In relation to the typologies discussed above, evaluation governance is thus governance supposed to work through evaluation in contrast to other types of governance such as bureaucracy which is supposed to work through rules.

SEG is one of two types of evaluation governance. The conventional form of evaluation governance is organisational evaluation governance. Organisational evaluation governance, sometimes also termed performance management or management by objectives (MBO), is anchored in evaluation criteria related to specific individual organisations and implemented in a hierarchy of a principal and one agent. In contrast, the new form of evaluation governance, the SEG, is implemented in organisational fields, such as policy sectors or fields of nation states, and aims at comparatively assessing several actors using a common set of evaluation criterion. Before focusing on different variants of SEG, the development of this type of evaluation practice will be briefly presented in a historical context.

Evaluation: The historical development

The practice of evaluation in the Nordic countries has developed across time in what have been described as waves. Since the 1960s, we have witnessed four waves, known as the scientific wave, the dialogue-oriented wave, the neo-liberal wave and the evidence-based wave (Vedung, 2010; Krogstrup, 2011). The waves are characterised as shifting, typical for the time, assumptions about the right approach to evaluation. It is however not clear how the waves have been consti-

tuted. Has the development been driven by methodological innovation and dialogue on which evaluation models to use or has it been driven by shifts in public sector policies defined as shifting views upon which governance types to promote?

According to my own experiences, another story about the historical development can be told if one look at which evaluands have been viewed as important across time. Thinking along these lines, three layers, all part of current evaluation practice, can be defined. The first layer, evaluation classic, initiated based on inspiration from the US, focuses on evaluation of programmes defined as experiments and public sector reforms. The idea here is that political initiatives and decisions should be followed up by evaluation activities focusing on implementation and effects. Following this idea, evaluation practice is organised ad hoc as a kind of follow-up research in the wake of important decisions and reforms.

The second layer, the above-mentioned performance management systems, developed based on inspiration from a larger set of Anglo-American countries, focuses on the evaluation of one agent. The idea is that in order to secure accountability and results, an organisation has to be monitored and evaluated by its principal. Following this idea, evaluation practice is organised as routines, typically as yearly recurrent processes defining goals and assessing goal attainment.

The third layer, the SEG, focuses on evaluation of a plurality of entities, organisations or nation states. The idea is that comparison and transparency in an organisational field facilitate learning and create a stimulating and problem solving environment for improving competition, dynamics sometimes discussed in the literature under the heading of naming and shaming (Pawson, 2002). Following this idea, evaluation practice is organised as recurrent routines comprising several organisations.

There are similarities, but also differences, between the first and the third layers, i.e., between evaluation classic and SEG. The similarity first and foremost is that both types of evaluation are enacted in organisational fields, the difference is that evaluation classic is enacted ad hoc while SEG incorporates recurrent routines. Below the different types of recurrent routines will be elaborated.

SEG: A typology

Fundamentally, we can distinguish between four different types of SEG: 1) Interorganisational SEG, 2) accreditation SEG, 3) systems- and process SEG and 4) evidence based SEG. Table 1 below presents the characteristics of the four SEG types on four dimensions. Two dimensions relate to the way governance is enacted, specifically referring to the principles upon which steering and transparency mechanisms are based. These are called the logic of steering and the logic of transparency. A third dimension, the logic of evaluation, relates to how evaluation is enacted. And a fourth, the logic of organisational behaviour, relates

to how organisational behaviour is understood. The four logics are elaborated below.

SEG may be viewed as administrative systems. Administrative systems can be anchored in different logics of steering. Drawing on Westerberg (2004; see also Forssell & Westberg, 2006: 191), we can distinguish between constituting and effectiveness promoting logics. Constituting steering logics aim at regulating the core activity of organisations, whereas effectiveness promoting steering logics aim at maintaining and developing organisations as such.¹

Table 1: A typology of different types of systemic evaluation governance.

| Logic: SEG: | Logic of steering | Logic of transparency | Logic of evaluation | Logic of organisational behaviour |
|---|--|----------------------------|--|--|
| Interorganisational SEG (benchmarking, ranking, rating, smileys etc.) | Constituting through making differences in performance visible. | Organisational and public. | Anchored in a result model measuring performance either on a scale or categorised. | Organisations are viewed as collective rational competitive actors aiming at performing as good as possible vis-a-vis competitors. |
| Accreditation SEG (programme- or institutional approval) | Constituting through third party control. | Organisational and public. | Anchored in a peer review or stakeholder model assessing whether specified thresholds are met. | Organisations are viewed as lazy, slack maximising actors who must be controlled to secure performance. |
| Systems- and process SEG (certification) | Effectiveness promoting through focus on systems- and process-oriented leadership. | Organisational. | Anchored in a system model. | Organisations are viewed as institutions able to secure performance through the development of quality assurance systems. |
| Evidence based SEG (systematic reviews, guidelines etc.) | Constituting through knowledge-based leadership. | Professional. | Anchored in an effect model assessing which interventions work. | Organisations are viewed as cognitive professional identities. |

SEG in addition may be viewed as transparency, or see-through, technologies, defined as systems designed to promote insight into organisations and organisational fields. The idea about transparency is complex and to some extent based on contradictory aspects. Transparency on the one hand may be seen as a democratic ideal. Citizens have the right to information concerning how resources generated through taxation are used. But transparency initiatives may also be seen as a tightening of monitoring and thereby as an expression of dis-

trust. Drawing on Levay (2006: 121) we can distinguish between three types of transparency logics. Transparency can be anchored in a organisational, public or professional logic.

An organisational transparency logic aims at seeing-through organisational performance in order to secure knowledge-based administration and leadership. A public transparency logic facilitates that all stakeholders, users, citizens, media as well as politicians are able to see-through performance. Public transparency may be related to user's free choice and thus aims to promote market competition, but it may also be related to the above-mentioned democratic ideal. Finally, a professional transparency logic aims at seeing-through professional practices in order to secure that professionals offer the best solutions to users and citizens.

Different types of SEG are anchored in different logics of evaluation. In the literature, typologies of evaluation models, defined as different organising principles for evaluation, are discussed (Hansen, 2005; Vedung, 1997, 2009). We can for example distinguish between result models, among these the goal attainment and the effect model, process models, system models, economic models, actors models, among these the client-oriented model, the stakeholder model and the peer review model.

Finally, different types of SEG are anchored in different perceptions of how organisations behave. Organisations for example may be viewed as rational collective actors aiming at solving the problems they are dealing with as effective as possible. Also, they can be viewed as lazy and slack, maximising actors aiming at cutting corners, as structural and cultural institutions or as cognitive professionals identities.

Interorganisational SEG

The performance of organisations in an organisational field can be assessed comparatively by using interorganisational SEG. Based on different types of indicators, organisations can be benchmarked and ranked or they can be rated using grades, for example based on numbers, stars or smiley faces.

Interorganisational SEG is anchored in a constituting logic of steering, as such, systems produce information about differences in the results of the organisations analysed. The logic of transparency is primarily organisational, as the information is supposed to be used for further developing organisational practices, but it may also be public, if there is an intention to influence the choice of potential users in relation to their choice of which organisational supply to demand. In an evaluation context, interorganisational SEG systems are anchored in result models. Results are assessed either on a scale or in categories. From an organisational perspective, interorganisational SEG views organisations as rational collective actors aiming at performing as well as possible by learning vis-à-vis their competitors.

Below, two examples of interorganisational SEG are presented. The first example, see box 1, concerns the field of employment policy. In this field, a

benchmarking portal has been developed. In the portal, the performance of public employment agencies (job centres) are compared. The aim is to facilitate a dialogue about the development in the field and inspire the job centres to develop their performance by creating a competition regarding results scores.

The National Labour Market Authority, which is part of the Ministry of Employment, has developed a web-based benchmarking tool that makes it easy to get an overview of labour market activities and the results of employability enhancement measures (see: www.jobindsats.dk – only in Danish). On the website, details on the number of employed and unemployed people as well as on the cost of efforts to get people into employment can be found. Information is available regarding the activities of the job centres as well as for municipalities, regions and on the national level. The job centres are compared in clusters facing similar labour market conditions.

The aim of the system is to facilitate a dialogue about the development in the field and inspire job centres to focus on how their performance may be developed. The job centres are able to examine themselves compared to other job centres on scorecards and use this knowledge to develop activities.

Box 1: Interorganisational SEG I: Employment policy

The second example, see box 2, concerns the field of health. Here, on the surface, a simple system gives hospitals stars for treatment quality. Looking deeper into the system, it is however obvious that the calculation methodology behind the stars is rather complex. The system in the field of health was developed following the introduction of better possibilities for patients' free choice between hospitals. The system is presented as a tool to support citizens/users as well as health professionals. The introduction of the system has raised a discussion about whether citizens/users are able to and wish to use the system or whether they prefer to use their doctor as the primary advisor. Hospitals getting good scores have been quick to demonstrate this in the public arena.

The two examples illustrate how logics of transparency differ in interorganisational SEG systems. In the example from employment policy, the logic of transparency is organisational, addressing management challenges. In the example from health policy, the logic of transparency is public, as the evaluative information aims at external stakeholders, citizens/users as well as professionals. This also shows that the two examples of logic of steering, which the two systems are expected to support, are different. In the employment example, leaders are expected to act on the information available. In the health example, users are expected to vote by their feet (exit a la Hirschman, 1970). In this case, market thinking is central.

Five years ago, the Ministry of Health launched a website called “health quality” (see www.sundhedskvalitet.dk – only in Danish). On the website, one can find hospitals’ grades, one grade for each hospital, and treatment grades, one grade for each treatment offered. The grades are presented as stars. A hospital can get from one to five stars, where five is the best. The system is made up of already existing information from several sources, including clinical data, as well as data on for example patient satisfaction. Grades are relative reflecting whether hospitals are worse or better than the national average.

The website is developed for the use of patients, relatives, hospitals and doctors. The aim is to support the free hospital choice of patients and their active withdrawal in own course of disease.

Box 2: Interorganisational SEG II: Quality in health care

Accreditation SEG

Accreditation is another type of SEG. Accreditation is an inspection method. It can be characterised as based on a traffic signal philosophy deciding on green light to go, red light not to go on or yellow light to go on under surveillance. The idea is that a third party approves, conditionally approves or disapproves a supplier basing approval or not approval on ex-ante specified evaluation criteria. Accreditation is anchored in a constituting logic of steering organised as a determination of whether practices meet minimum standards. The logic of transparency is organisational and public. The logic of evaluation draw on a peer review or a broader stakeholder model. The logic of organisational behaviour is anchored in a view of organisations being lazy. Because organisations are lazy, inspection is needed to assure quality.

As the two examples in the following show, accreditation may focus on programmes or organisations and accreditation may be compulsory or voluntarily.

The Danish accreditation model in the higher educational field, see box 3, is compulsory. All educational programmes have to meet specified evaluation criteria. Some higher educational institutions, especially in the field of business economics, also participate in voluntarily international systems of accreditation. In this context, accreditation is a kind of ticket to a special “society”. Accreditation is a quality mark experienced as necessary to compete in the market. Formally it is voluntarily, but as it is experienced as a prerequisite for organisational success or maybe even survival, it can be interpreted as voluntariness under constraint.

The Danish accreditation model in healthcare, see box 4, is compulsory for hospitals, public as well as private, but voluntarily for pharmacies and municipalities. In the field of pharmacy, nearly all pharmacies however participate. It seems to be difficult to not participate.

The two examples illustrate that accreditation also can be characterised by different logics of transparency. In the higher education example, the logic of transparency is primarily public, as the evaluative information is passed on to students and the labour market, believing that this will develop a well-functioning educational market. In the healthcare example, the logic of transparency is mainly organisational. Here the aim is to promote collaboration and quality development. The logics of steering in the two examples also differ. The market is seen as the driver in the education example, whereas the driver in the healthcare example is organisational accountability promoted by external inspection.

As part of the Bologna process, the European ministers of education in 2005 decided on a common set of standards and guidelines for quality assurance of higher education. The standards concern three themes: 1) Internal quality assurance at higher educational institutions, 2) external quality assurance of educational programmes or institutions and 3) the organisation of external quality assurance authorities. The standards define the frames for national quality assurance practices but do not define a special model for these.

In Denmark, the Parliament in 2007 decided that the Danish quality assurance model should be based on accreditation of educational programmes. Every programme existing, as well as new ones, should pass an accreditation process. This means that around 1000 programmes at the universities have to be accredited. Every one of these has to be able to document that it meets specified minimum standards related to the societal relevance and demand for the programme, that the programme is anchored in research, that there is consistency between the programme's aims concerning learning outcomes, its title and content, that it is organised appropriately and that there are routines for internal quality assurance.

The aim of the accreditation system is to contribute to creating a more coherent and transparent educational market to the benefit of students, the labour market and the educational institutions.

Box 3: Accreditation SEG I: Higher education

In the mid 1990s, the idea of developing a Danish quality model for health care was initiated. In 2005, the Danish Institute for Quality and Accreditation in Healthcare (IKAS, se www.ikas.dk) was established. The Danish Healthcare Quality Program (DDKM) aims at promoting collaboration in the healthcare system, creating better and more coherent patient care and continuous quality development as well as documenting and making visible the quality in the healthcare system. DDKM is an accreditation system. In each sector and field, today hospitals and pharmacies, but in the long run also municipalities and probably also both general and special practitioners, specified accreditation standards are worked out.

The standards are comprehensive. The first version for the hospitals was made up of 104 standards. There are standards for every possible aspect, for example leadership, quality and risk management, documentation, hiring and competence development, hygiene, recovery plans in case of emergency, equipment, patient involvement, patient information, admission ward, planning, diagnosis, medication, observation, different types of treatment as well as nutrition and prevention. Add to this a long list of accreditation standards for treatment of specific diseases. In the field of pharmacies, the first version is made up of 42 accreditation standards.

As part of the accreditation process, organisations create a self-evaluation of whether they meet standards. After this, they are visited by impartial professionals assessing the case. Finally, an accreditation council decides on the status of accreditation. There are three possible decisions. If an organisation meets the standards, it receives a certificate for three years, whereafter the process starts once more. If an organisation receives critical comments and only partly meet the standards, it subsequently has to document that things not approved at first are solved. If an organisation is not approved, it gets a deadline for solving things, whereafter it is revisited by the professionals. If there are still problems, it gets the status as non-accredited. Among the hospitals which have gone through the process, most have been accredited with critical comments. Only a few have gone through the process without receiving some critical comments. None to this point have received the status as non-accredited.

Box 4: Accreditation SEG II: Health care

Systems and process SEG

Systems and process SEG is anchored in a more indirect approach to development of performance than the above analysed SEG types. Systems and process SEG aims at developing a local evaluation culture. Systems and process SEG is

based on an effectiveness promoting logic of steering as requirements are related to quality assurance system capacity. The logic of transparency is organisational and the logic of evaluation based in a system model. The organisational logic is that if demands to develop structure and culture are met, there is no need for direct control of organisational results.

As the term indicates, there are two variants of systems and process SEG: certification and process requirements.

Certification is like accreditation, a traffic signal philosophy, but in certification, the demand is for the organisations to establish quality control systems, which are able to detect and solve quality problems. Organisational systems and procedures, and not the core production of organisations, are the focus. As in accreditation, a third party is responsible for the approval process, which may be compulsory or voluntary.

In Box 5, an example from the environmental area is presented. Another example is the Norwegian model for meeting the European standards in higher education. In Norway, these are implemented mainly through certification (in Norway termed audit) and not as in Denmark through a rigid accreditation system.

In 2007, a comprehensive structural public sector reform was introduced in Denmark. Municipalities were merged, counties were abolished and new and larger regions were established. Responsibility for tasks was moved across public sector levels. The municipalities were given responsibility for several tasks related to natural resource management and the environment formerly placed in the old counties. There were severe doubts whether municipalities would be able to maintain the needed specialised expertise, as task responsibility would be spread to a much larger number of organisational entities. This criticism made Parliament decide that every municipality should establish a quality assurance system concerning their regulatory processes and that these systems should pass certification. The systems developed in the municipalities concern procedures for case work as well as for the handling of documents, implementation of user satisfaction evaluations and procedures for an early internal evaluation of the quality assurance system. It is obvious that this kind of system has the potential for assuring quality in the organisation of case work, but this does not necessarily imply that the professional regulatory decisions have quality. The policy decision in Parliament to adopt a quality assurance system seems to be more symbolic than actually meeting the criticism of the reform.

Box 5: System and process SEG I: Certification in the environmental field

The certification system in local government in the field of natural resource management and environment is compulsory, but there is no established special-

ised organisational approval entity, as in the accreditation system of higher education. Instead, the municipalities can choose between suppliers in the market for certification. Also, there are no sanctions linked to the system and there are examples of municipalities whose certificates have been suspended for periods of time due to lack of resources.

Another example can be found in the field of public schools, see box 6. Here it is compulsory for municipalities to create the so-called quality reports. Also, the quality reports are based on an effectiveness promoting logic of steering. The idea is to force the municipalities through a reflection process, specifying how they work to assure and develop the quality of public schools. As this system also demands specification of how the municipalities work to develop quality on selected substantial dimensions, the system also has some elements of a constituting steering logic.

In 2006, the Parliament decided to change the law on public schools. The aim was to strengthen the evaluation culture. National tests and individual plans for pupils were introduced and it was decided that the municipalities should work out a yearly quality report. The report has to describe the school system, the quality level of the individual schools, how the municipality works with quality development and what has been done since the last quality report was prepared. If a school does not meet a satisfactory quality level, an action plan has to be created. Quality reports and action plans have to be sent to school boards for comments and published on the Internet.

The quality reports are planned to be the quality assurance system of the municipalities and they help hold municipalities accountable for quality. In addition, the system promotes dialogue and collaboration about evaluation and quality development among actors at different levels and gives parents quality information.

Some minimum requirements for the content of quality reports are defined, but the room for manoeuvring for the municipalities is considerable. The National Agency for Quality and Supervision, which is part of the Ministry of Children and Education, monitors the system and advises municipalities in questions concerning how to work with quality reports.

Box 6: System and process SEG II: Quality reports in the field of public schools

Evidence based SEG

The concept of evidence has drawn attention in recent years. The idea is that policies and practices, i.e., for doctors, nurses, teachers and leaders, should be

evidence-based, defined as anchored in the best possible knowledge about which interventions work.

Krogstrup (2011) describes this development as a pendulum swinging back to the classical thinking about the effect model in the evaluation literature. This is correct in the way that current evidence advocates build upon this tradition. But at the same time, it is a truth with modifications. Firstly, there is an ongoing discussion on whether to define evidence narrowly as effect information produced by randomised trials or define evidence broadly as information also produced by collecting professional experiences and client-oriented knowledge? Secondly, compared to previously, the current evidence movement is organised much more systematically than was the case earlier. New specialised evidence producing organisations have been established. Some of these are international, such as for example Cochrane Collaboration in the medical field and Campbell Collaboration, working in the fields of social, educational and criminology policies, others are national, such as for example the Danish Clearinghouse in the educational field (Hansen & Rieper, 2009, 2010; Rieper & Hansen, 2007).

Common to the evidence producing organisations is that they build up, update and spread research and knowledge reviews. Evidence thinking is anchored in a constituting logic of steering, as specific demands on interventions are formulated. The evidence movement spreads knowledge which professionals may adopt and as such is mainly anchored in a professional logic of transparency.

One can pose the question whether evidence thinking also is anchored in a public user oriented logic of transparency. The spontaneous answer is yes. Users are able to download information about which intervention works the best given specified diagnoses and problems. On this basis, they will be able to confront visitation officers and professionals with this knowledge in order to get the best possible individual treatment, or through patient associations they may be able to confront politicians. However, in reality the answer is not that simple. Research and knowledge reviews are most often very specialised reports and very hard to read and understand for non-professionals. Maybe we can speak about evidence thinking as based in a public super-user oriented logic of transparency?

Heading towards still more complex SEG?

Above I have argued that a new type of evaluation termed SEG has been developed and increasingly spread in recent years. Also I have developed a typology of the different forms of SEG and illustrated this with concrete examples. The examples have shown that some SEGs are characterised by practices combining elements from several of the forms characterised in the typology.

Recently, the Danish minister responsible for the public schools presented an action plan for strengthening bilingual pupils' competences (Undervisningsministeriet, 2011). The action plan proposes to establish SEG in still another policy area. The proposal is interesting, as this SEG if established will combine elements from still more forms of the SEG typology compared to the examples presented above.

The SEG for strengthening bilingual pupils' competencies, see box 7, combines evaluation, intensified ministerial control and evidence thinking. The aim is to use effective evaluation of interventions to spread knowledge about what works across schools.

Analyses, among others the so-called PISA Ethnic 2009, have shown that many bilingual pupils leave school without having basic reading competencies. They also show that there are differences across schools. Reading competencies are weak in schools with high shares of immigrant pupils (Undervisningsministeriet, 2011). With this background, the minister has proposed to carry through an early analysis of bilingual pupils' reading competencies. The analyses shall combine existing data on pupils' grades, proportion of pupils passing on to post-secondary education, the results of national tests administered to all pupils at all schools in several subjects, as well as the so-called educational preparedness assessments. All this information on pupils' competencies will be coupled to school information, for example regarding the proportion of bilingual pupils, the attendance of pupils and teachers and the teaching frequency of teachers.

The idea is that schools and municipal school administrations shall use the analyses to monitor the development of local results compared to the results of the rest of the country in order to decide whether special initiatives are needed. Also, the analyses shall help politicians at national and local levels in their control activities. Schools which steadily are among the 5% worst performing will be met with intensified control and dialogue about how to solve problems. The SEG will also include effect evaluation of interventions and pedagogical methods at schools with high shares of bilingual pupils and good results. The aim of this is to secure that schools are able to learn from each other.

Box 7: Action plan for bilingual pupils competencies

The steering logic of the proposed system is constituting. The transparency logic is organisational, public and professional. The evaluative information produced is meant for school leaders, school administrations and politicians, whose behaviour is expected to be influenced if the evaluative information turns on warning lights about the results. The SEG in other words is meant to be a lever for managerial development initiatives. If these not triggered by the institutional competition logic and the evidence-based component, the SEG is also constructed with a possibility to intensify control. This SEG is interesting also due to its task selectivity, as it focuses on a special dimension in the total task profile of public schools. The proposal may be interpreted as yet another national govern-

ance initiative, initiated by the central government for public schools at local governmental level.

Conclusions

This article has analysed SEG as a modern field-based governance practice aiming at challenging and developing public organisations. In the introduction, two research questions were raised. These concerned the characteristics of SEG conceptually and as practice, as well as the logics of different forms of SEG. The questions have been answered by defining SEG conceptually and by developing and illustrating a typology of SEGs. As opposed to more conventional forms of evaluation, such as ad hoc programme evaluation and performance management routines in a principal-agent relation, SEG is routine-based and comparative. SEG works through spreading evaluative information in organisational fields, aiming at increasing learning and competition across actors and organisations.

As shown, there are four types of SEG. Interorganisational SEG systems build up evaluative information upon which actors are expected to act constructively to develop their performance. Accreditation systems also build up evaluative information, but here third parties approve whether organisations meet in advance decided minimum performance standards. System and process SEG systems have a more indirect approach to performance. Here, the idea is to develop and sometimes also approve the evaluation culture of organisations. Finally, evidence-based SEG systems build up and spread evaluative information on interventions, focusing on whether interventions work. Organisations are expected to dismantle interventions with no evidence for effects and adopt interventions with evidence for effects.

The overall purpose of the article has been to develop an analytical frame of reference suitable for more thorough empirical analyses. The examples have been presented based on documentary material alone. The consequence of this is that the pictures drawn are formal and expressions of how SEGs are legitimised by their initiators. One could also argue that the article has addressed SEGs at their surface levels although unfolding logics their initiators not always make explicit. A crucial question is which faces can be found beneath the SEG surfaces? Does SEG, as expected, facilitate organisational learning? Does the competition element of SEG, as expected, further develop organisational performance? Are there other dynamics and effects, such as, for example, de-coupling of SEGs from organisational practices or standardisation of organisational practices? Are SEGs (in line with the argument of Rose, 1999) advanced surveillance systems for correcting and normalizing professionals and public organisations? A Danish study of benchmarking of hospitals indicates this (Triantafyllou, 2007). In order to uncover whether this also applies to other types of SEG more thorough analysis of the implementation, enactment and effects of SEGs based on other types of data than the ones used here, such as interviews, surveys and observations of dynamics across time, should be carried out.

References

- Forssell, Anders & Anders Ivarsson Westerberg (2006) "Vad kostar transparens?" in Charlotta Levay & Caroline Waks (eds.): *Strävan efter transparens*. SNS Förlag, Stockholm.
- Fournier, Deborah M. (2005) "Evaluation" in Sandra Mathison (ed.) *Encyclopedia of Evaluation*. Sage Publications, Thousands Oaks.
- Hansen, Hanne Foss (2005) "Choosing Evaluation Models. A Discussion on Evaluation Design." *Evaluation*, vol. 11 (4): 447-462.
- Hansen, Hanne Foss & Olaf Rieper (2009) "The Evidence Movement. The Development and Consequences of Methodologies in Review Practice." *Evaluation*, vol. 15 (2): 141-163.
- Hansen, Hanne Foss & Olaf Rieper (2010) "Institutionalization of Second-Order Evidence-Producing Organizations" in Frans L. Leeuw, Tom Ling & Olaf Rieper (eds.) *The Evidence Book. Concepts, Generation, and Use of Evidence*. Transaction Publishers, New Brunswick.
- Hirschman, Albert O. (1970) *Exit, Voice and Loyalty: Responses to Decline in Firms, Organizations and States*. Harvard University Press, Cambridge MA.
- Howlett, Michael (2009) "Governance modes, policy regimes and operational plans: A multi-level nested model of policy instrument choice and policy design." *Policy Sciences* 42, 73-89.
- Jørgensen, Torben Beck & Bøje Larsen (1982) "Styring – et forsøg på teoridanelse" in F. Agersnap m.fl. *Forskningsstyring og forskningsmiljø*. Nyt fra Samfundsvidenskaberne, København.
- Krogstrup, Hanne Katrine (2011) *Kampen om evidens. Resultatmåling, effekt-evaluering og evidens*. Hans Reitzels Forlag, København.
- Levay, Charlotta (2006) "De nationale kvalitetsregistren som verktøj for transparens" in Charlotta Levay & Caroline Waks (red.): *Strävan efter transparens*. SNS Förlag, Stockholm.
- Pawson, Ray (2002) "Evidence and Policy and Naming and Shaming." *Policy Studies* 23 (3/4), 211-230.
- Pierre, Jon (2000) *Debating Governance: Authority, Steering, and Democracy*. Oxford University Press, Oxford.
- Rieper, Olaf & Hanne Foss Hansen (2007) *Metodedebatten om evidens*. AKF Rapport, København.
- Rose, Nikolas (1999) *Powers of Freedom*. Cambridge University Press, Cambridge.
- Triantafillou, Peter (2007) "Benchmarking in the Public Sector: A Critical Conceptual Framework." *Public Administration* 85 (3), 829-846.
- Undervisningsministeriet (2011) *National handlingsplan for styrkelse af tosprogede elevers faglighed*. København.
- Vedung, Evert (1997) *Public Policy and Program Evaluation*. Transaction Publishers, New Brunswick.

- Vedung, Evert (1998) "Policy instruments: Typologies and Theories" in Bemelmans-Videc, Marie-Louise, Ray C. Rist & Evert Vedung (eds.) *Carrots, Sticks and Sermons. Policy instruments and their evaluation*. Transaction Publishers, New Brunswick.
- Vedung, Evert (2009) *Utvärdering i politik och förvaltning*. Studentlitteratur, Lund.
- Vedung, Evert (2010) "Four Waves of Evaluation Diffusion." *Evaluation* 16 (3), 263-277.
- Westberg, Anders Ivarsson (2004) *Papperspolisen. Den ökande administrationen i moderna organisationer*. EFI: Handelshögskolan i Stockholm.

Notes

¹ Westerberg (2004) operates with a third type based on a legitimating logic. In some cases, implementation of SEG may aim mainly at legitimating an organisational field. This aspect is not included in the analysis here, as the research question raised concerns the more functional assumptions of different types of SEG.